On the Design of Generalist Strategies for Swarms of Simulated Robots Engaged in Task-allocation Scenarios

Tuci, Elio; Rabérin, Alexandre

Published in:
Swarm Intelligence

DOI:
10.1007/s11721-015-0113-y

Publication date:
2015

Citation for published version (APA):
Tuci, E., & Rabérin, A. (2015). On the Design of Generalist Strategies for Swarms of Simulated Robots Engaged in Task-allocation Scenarios. *Swarm Intelligence*, *9*(4), 267-290. https://doi.org/10.1007/s11721-015-0113-y

Swarm Intelligence manuscript No.
(will be inserted by the editor)

On the Design of Generalist Strategies for Swarms of Simulated Robots Engaged in a Task-allocation Scenario

Elio Tuci · Alexandre Rabérin

Received: date / Accepted: date

Abstract This study focuses on issues related to the evolutionary design of task-allocation mechanisms for swarm robotics systems with agents potentially capable of performing different tasks. Task allocation in swarm robotics refers to a process that results in the distribution of robots to different concurrent tasks without any central or hierarchical control. In this paper, we investigate a scenario with two concurrent tasks (i.e., foraging and nest-patrolling) and two environments in which the task priorities vary. We are interested in generating successful groups made of behaviourally plastic agents (i.e., agents that are capable of carrying out different tasks in different environmental conditions), which could adapt their task preferences to those of their group mates as well as to the environmental conditions. We compare the results of three different evolutionary design approaches, which differ in terms of the agents' genetic relatedness (i.e., groups of clones and groups of unrelated individuals), and/or the selection criteria used to create new populations (i.e., single and multi-objective evolutionary optimisation algorithms). We show results indicating that the evolutionary approach based on the use of genetically unrelated individuals in combination with a multi-objective evolutionary optimisation algorithm has a better success rate then an evolutionary approach based on the use of genetically related agents. Moreover, the multi-objective approach, when compared to a single objective approach and genetically unrelated individual, significantly limits the tendency towards task specialisation by favouring the

Elio Tuci
Llandinam Building, Computer Science Department
Aberystwyth University
Aberystwyth, SY23 3DB
UK
Tel.: +44-(0)1970-622537
Fax: +44-(0)1970-628536
E-mail: elt7@aber.ac.uk

Alexandre Rabérin
Llandinam Building, Computer Science Department
Aberystwyth University
Aberystwyth, SY23 3DB
UK
E-mail: raberin.alexandre@gmail.com

emergence of generalist agents without introducing extra computational costs. The significance of this result is discussed in view of the relationship between individual behavioural skills and swarm effectiveness.

Keywords Task-allocation · Evolutionary Swarm Robotics · Multi-objective optimisation algorithms

1 Introduction

In natural swarms, colonies of ants and bees manage to efficiently perform vital concurrent tasks, such as foraging, defence, nest construction and maintenance, by balancing the resources on different tasks according to the current colony needs. Workers are engaged in specific tasks through a task-allocation process that operates without any central control (Page and Mitchell, 1998). Age and morphological characteristics of the workers have been initially considered the main factors that influence the task-allocation process. However, few studies have shown that task allocation in social insects can be guided by other emergent circumstances concerning the life of the colony (Gordon, 1996; Page, 1997; Gordon, 1989). For example, in (Seeley, 1989) it is shown that a honeybee forager's decision on whether to collect nectar or remain in the nest depends on how much nectar is already stored in the nest.

Similar to social insects, swarm robotics systems are a multi-robot systems that develop self-organised collective responses without central control using local communication strategies (Dorigo and Şahin, 2004; Dorigo et al., 2014). The interest in swarm robotic systems is often related to what these systems could offer in terms of automating parallel processes requiring collection and transport of materials (e.g., toxic waste), or the assembly of structures using basic building blocks (Dorigo et al., 2004; Allwright et al., 2014). These challenging scenarios in which, for reasons of efficiency, the robots of a swarm have to distribute themselves on different tasks, require task allocation, a process that results in specific robots being engaged in specific tasks. During the task-allocation process in a swarm robotics system, the robots select particular tasks without any central or hierarchical control. In this study, we look at task allocation from the perspective of Evolutionary Swarm Robotics (hereafter, ESR), where the robot behavioural mechanisms are automatically generated using evolutionary computation techniques to synthesise artificial neural network controllers (Trianni and Nolfi, 2011). ESR helps the designer to circumvent the problem of decomposing the group response in individual behaviours and underlying mechanisms, by generating both through an autonomous process inspired by natural evolution (Nolfi and Floreano, 2001). ESR methods can be used to design controllers for robot swarms operating in scenarios in which the number of individuals engaged in each task changes according to the current swarm needs, and in which task allocation emerges from the interactions of individuals that simply react to contingent events, such as changes in environmental conditions, robot failure, etc. (Tuci and Trianni, 2012).

In this study, we use the term *behavioural plasticity*, in the context of a task-allocation scenario for swarm robotics systems, to refer to the capability of a single robot to carry out different tasks in different environmental conditions. A robot lacks behavioural plasticity if it can only execute a single task regardless of the

environmental conditions. Such a robot can also be referred to as a *specialist agent*. On the other hand, a robot is considered fully behaviourally plastic if it exhibits the ability to carry out all tasks in all environmental conditions. Such a robot can also be referred to as *generalist agent*. Specialist and generalist are terms that define the extreme points of the behavioural plasticity spectrum, which will be operationally defined in later Sections of this study.

For a swarm of robots operating in a scenario requiring task allocation, behavioural plasticity is an important property of single agents. Generalist agents can allow a swarm to successfully face different potentially disruptive phenomena by re-distributing resources according to current swarm needs. For example, generalist agents can allow a swarm: i) to adjust to changes in the task priority; ii) to cope with conditions in which tasks are distributed over time, and in which at different times of the swarm life different sub-sets of tasks need to be carried out; and iii) to cope with variations within the swarm, caused either by the replacement of agents with other agents, or by changes to the cardinality of the swarm. For instance, a swarm can increase or decrease in terms of number of agents, with consequent variations in the number of generalist and specialist agents within the swarm.

Designing generalist agents for a swarm of robots that can face all the sources of variation mentioned above using evolutionary methods is a still unsolved scientific challenge. Our goal is to move a step forward in the development of effective design methods that could overcome some of the current limitations of the evolutionary approach when applied to swarm robotics. In this study, we focus exclusively on circumstances in which the behavioural plasticity is required by the agents to cope with variations within the swarm. In other words, our objective is to design controllers for generalist agents that, in the context of a specific task-allocation scenario, can adjust their task preferences according to the preferences of their group mates. The primary contribution of this study is to demonstrate that this type of behavioural plasticity can be obtained by using a relatively efficient evolutionary method that, contrary to other solutions, helps the designer to overcome several undesired effects. This method is based on the use of swarms of genetically unrelated agents in combination with an evolutionary multi-objective optimisation approach in which the selective advantage of swarms with generalist agents over swarms without generalist agents is simply determined by a fitness objective.

In the next Section, we begin with the concept of genetically related and unrelated agents. We then illustrate the principles of the design approaches used in ESR. With this background, we continue by illustrating our motivations, goals, and achievements, starting from a brief description of the work published in (Quinn, 2001), which effectively represents the source of inspiration for our research study.

2 Background and Motivations

Generally speaking, there are two main approaches to the evolutionary design of controllers for a swarm of robots: the *clonal approach* and the *aclonal approach*. The two approaches mainly differ in the way groups are formed during evolution. In the clonal approach, the groups are homogeneous, since individual controllers of each group of robots are formed from a single genotype within one population of genotypes. Robots are clones. In the aclonal approach, groups are heterogeneous since

individual controllers of each group are formed from multiple genotypes within one population of genotypes. Robots are not clones.

In a swarm robotics scenario requiring task allocation, the robots of a successful homogeneous group generated with a clonal approach are behaviourally plastic. Multiple factors including the evolutionary conditions and nature of the neural mechanisms determine whether, and for how long, individual plasticity can be retained. However, at least of the beginning of a trial, all robots are the same, and all of them can potentially perform the tasks outright. Thus, the clonal approach appears to be the best option to generate plastic individuals for swarms capable of (re)distributing resources to concurrent tasks according to the swarm needs. Still, data show that the clonal approach is not as effective as the aclonal approach in finding groups that successfully solve task-allocation scenarios (Quinn, 2001; Tuci, 2014). As originally shown in (Quinn, 2001), the advantage of the aclonal approach over the clonal approach can be accounted for with reference to specialisation[1]. In the clonal approach, the evolutionary path to the emergence of successful groups is limited to only those trajectories in which the mechanisms for allocating the robots to tasks, and the mechanisms for executing the tasks, originate at the same time (Quinn, 2001; Tuci and Trianni, 2014). This is because the emergence of either the former or the latter set of mechanisms, in the absence of the complementary part, is neither beneficial to individuals nor to groups. Thus, solutions take generally longer to evolve (i.e., more generations) and are more difficult to find (i.e., less evolutionary runs generate successful groups).

Contrary to the clonal approach, the aclonal approach can generate groups of minimally plastic or specialist agents in which the allocation of tasks is simply based on individual competencies. No complex task-allocation mechanisms are required to distribute robots to tasks. Thus, successful groups are generally found quicker (i.e., in less generations) and more often (i.e., more evolutionary runs generate successful groups). However, the adaptability of these groups, with respect to environmental variations, is significantly restricted by the limited behavioural competencies of specialist agents. For example, heterogeneous groups generated with the aclonal approach fail to cope with environmental conditions in which a task requires more agents than those specialised on that task.

From an evolutionary design perspective, it would be extremely advantageous if we could exploit the benefits of the aclonal approach, and, at the same time, find ways to generate groups in which at least some of the individuals are behaviourally plastic. Quinn (2001) explored this issue in an extremely simplified scenario, where two robots are required to decide whether to take the leader or follower role in a task requiring coordinated movement. The roles were allocated only once and never renegotiated during evaluation. In spite of the simplicity, the study shows that even with the aclonal approach it is possible to generate successful groups of behaviourally plastic, instead of specialist, agents. Tuci (2014) compared the clonal and aclonal approaches in a more challenging scenario featuring two tasks and two operating (environmental) conditions, with swarms that have to allocate more resources (i.e., the majority of robots) to one task or the other according to the current environmental condition. This implies that at least some of the robots of a group have to switch tasks to balance the resources according to the "rules

[1] A more comprehensive discussion of the advantages of the aclonal over the clonal approach can be found in (Tuci and Trianni, 2014).

of the game". In this type of scenario, individual behavioural plasticity can be measured with reference to the frequency with which a robot performs each task in each type of environmental condition. The results of this study partially confirm what was illustrated in (Quinn, 2001), demonstrating that the aclonal approach outperforms the clonal approach in generating successful swarms. Nevertheless, no evidence of behavioural plasticity in aclonally generated swarms was shown.

In this study, we intend to deepen the exploration into the potentialities of aclonal approaches for the design of robotic swarms that successfully operate in scenarios requiring task allocation and continuous re-distribution of robots to tasks. Borrowing the task originally developed in (Tuci and Trianni, 2012), we run a comparative analysis in which robots' controller are synthesised using a clonal approach and two different types of aclonal approaches. In order to facilitate the emergence of behaviourally plastic agents in aclonal approaches, we developed an evaluation function made of two parts: the first part targets the quality of the swarm performances in solving the group task; the second part targets the individual plasticity of the agents of a swarm. In our study, the two aclonal approaches differ in terms of the characteristics of the evolutionary algorithm used to evolve heterogeneous groups. In the *evo-aclonal-single* approach, we use a single-objective evolutionary algorithm in which groups are scored according to an evaluation function that takes into account, in an additive way, the two fitness components above mentioned. In the *evo-aclonal-multi* approach, we use a multi-objective evolutionary algorithm in which each of the two fitness components above mentioned represents a different objective. We chose to operate at the level of the fitness function because steering evolution toward the emergence of swarms with generalist agents through the addition of a fitness objective is a computationally efficient approach compared to alternative approaches. For example, selective pressures favouring swarms of generalist agents in aclonal evolution can be implemented by evaluating the swarms' performance on tasks that can only be solved by swarms with generalist agents. Alternatively, the evaluation of agents in differently assorted groups can also favour the emergence of the mechanisms underpinning behavioural plasticity. Unfortunately, most of the solutions aiming at the development of behavioural plasticity through task requirements or through variations of the evaluation criteria tend to require longer evaluation times and consequently they increase the computational costs of the design algorithm.

The results of this study clearly show that the advantage of the aclonal over clonal approaches discussed above can be exploited without sacrificing the agents' behavioural plasticity, if the aclonal approach is used in combination with an evolutionary multi-objective optimisation algorithm. We show that the *evo-aclonal-multi* approach (i.e., the aclonal evolutionary multi-objective optimisation algorithm) is an efficient way to design controllers for successful robotic swarms operating in a task-allocation scenario. We also show that the *evo-aclonal-multi* approach allows the designer to retain into the evolving populations a larger number of behaviourally plastic agents compared to the *evo-aclonal-single* approach. Post-evaluation analyses shed light on the limits of the *evo-aclonal-single* approach, highlighting the benefits of the *evo-aclonal-multi* approach in generating behaviourally plastic agents. This work contributes to the development of more effective ESR design methods by exposing the benefits of an evolutionary approach for the development of behaviourally plastic agents for swarms of robots engaged in task-allocation scenarios. Further implications of the results of this study and reference

to relevant literature are discussed in Section 8. In Section 3, 4, 5, 6 we describe the task and the design methods. In Section 7, we illustrate the results of the evolutionary runs and of post-evaluation tests.

3 The Task

Groups comprising five simulated e-puck robots are evaluated in a scenario requiring task allocation. Hereafter, the term *simulated robot*, and *agent* are used to refer to the simulated model of the e-puck robot, detailed in Section 4. Taking inspiration from the behaviour of social insects, the two tasks to be performed by the simulated robots are called nest patrolling and foraging (hereafter, we refer to them as *task P*, and *task F*, respectively). Roughly speaking, *task P* requires an agent to remain within the nest. *Task F* requires an agent to leave the nest for the foraging site, to spend a certain amount of time at the foraging site, and then to come back to the nest. A group is required to execute both tasks simultaneously. Therefore, the agents have to go through a task-allocation phase in which they autonomously decide who is doing what, and then execute their respective tasks[2]. Moreover, the simulated robots are required to be able to switch from one task to the other due to the fact that they experience two different types of environment, *Env. A* and *Env. B*. In *Env. A*, *task F* is more important than *task P*. This means that in *Env. A*, a group maximises the fitness if the majority of agents (i.e., more than two agents) visits the foraging site and the minority (i.e., less than three agents) remains in the nest. In *Env. B*, *task P* is more important than *task F*. This means that a group maximises the fitness if the majority of agents remains in the nest and the minority visits the foraging site. Since a group, throughout its life-span, experiences both types of environment, not all the agents can specialise on a single task. At least one agent has to be able to play both tasks and eventually to switch from one task to the other based on the current environmental condition and the tasks allocated to the other group mates.

A group is required to keep track of the environmental changes with an effective (re)distribution of agents to tasks. As detailed below, in aclonal approaches, during evolution the frequency with which a simulated robot performs each task in each type of environment will be used as a measure of an agent behavioural plasticity. With the term *specialist agent* we indicate a simulated robot that tends to have a strong preference for a specific task regardless of the environment; while a *generalist agent* is a simulated robot that tends to execute with roughly equal frequency all the four task-environment combinations.

Note, this task does not necessarily require highly plastic agents for a group to be successful. An optimal performance can be achieved by a group in which two agents are specialised in *task P*, two agents are specialised on *task F*, and one agent is minimally plastic to be able to perform both tasks but not necessarily both of them in *Env. A* and in *Env. B*. The reader may wonder why would anyone be interested in evolving behaviourally plastic agents for a task-allocation scenario that does not require this type of competence. The answer can be found in the computational overhead that scenarios requiring generalist agents tend to produce

[2] Note that this is just a linguistic description of the task-allocation process required by this scenario. This description should not be interpreted as an operational illustration of the agents' behaviour.

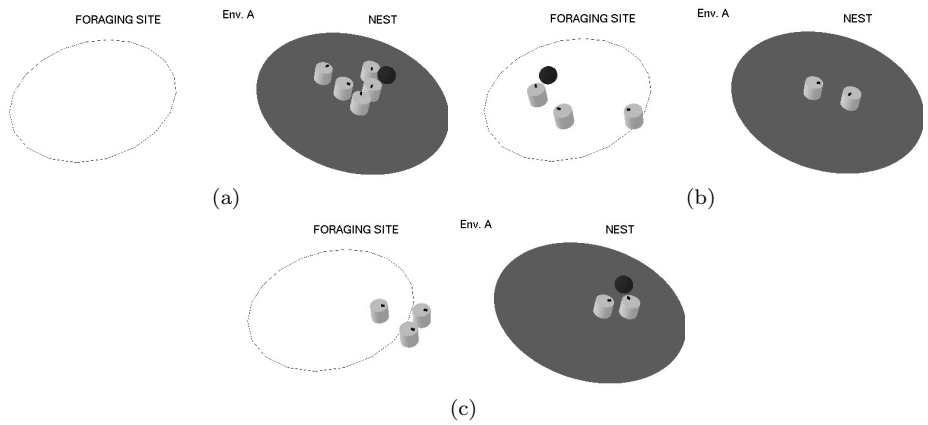

Fig. 1 Experimental scenario. Simulator snapshots taken in *Env. A* during: a) *Phase 1*; b) *Phase 2*; c) *Phase 3*. In each figure, the dark grey filled circle is the nest; the light grey cylindrical objects are the simulated robots; the black filled circle is the light. For visualisation purposes, the black circumference delimits the foraging site.

in the context of ESR. Many real world applications for swarms of robots can be undoubtedly complex and generally set in environments where the operating conditions tend to change over time. Behavioural plasticity is consequently a desirable characteristic of the agents that provides swarms the required flexibility to cope with potentially disruptive events (e.g., robots failures, changes in task priority, changes in the swarm composition, etc.). Nevertheless, from an evolutionary design approach, the idea of steering evolution towards, or generating the selective pressures for, the emergence of behavioural plasticity by modelling these potentially disruptive events is computationally costly. This is because modelling the above mentioned events tends to demand a larger evaluation time, and consequently more time to complete the evolutionary process. Our objective is to develop alternative evolutionary methods which facilitate the evolution of behaviourally plastic agents for swarms of robots required to operate in task-allocation scenarios. We aim to achieve this without having to pay the computational costs required to model, during the design phase, the events that make behavioural plasticity an adaptive trait. Specifically, we intend to exploit the advantages offered by the use of heterogeneous groups and the benefits offered by the use of an evolutionary multi-objective optimisation approach in order to design effective swarms with behaviourally plastic agents.

4 The Simulation Environment

In the foraging scenario studied in this paper, the environment is a boundless arena with a nest and a foraging site. The nest is a circular area in which the colour of the floor is in shades of grey (see Figure 1a). The radius of the nest is randomly defined at the beginning of each trial in the interval [20 cm, 30 cm]. The colour of the floor in the nest is dark grey in *Env. A*, and light grey in *Env. B*. The intensity of grey colouring the floor in the nest site is the only means the agents have to

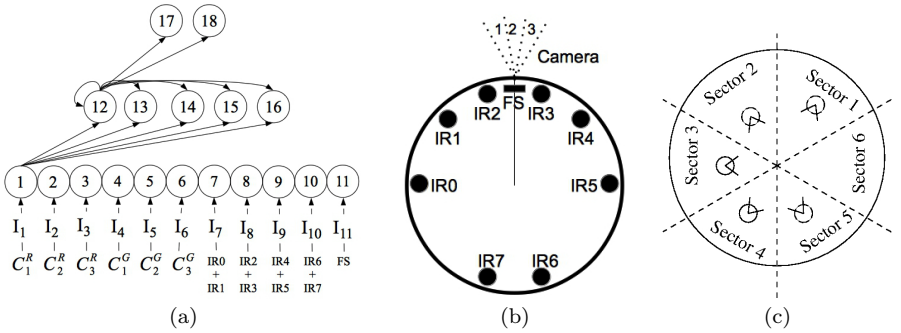

Fig. 2 (a) The neural network. Continuous line arrows indicate the efferent connections of the first neuron of each layer. Neurons on the same layer share the same type of efferent connections. Underneath the input layer, it is shown the correspondences between sensors and sensor neurons. (b) E-puck body-plan. The black circles refer to the position of the infra-red (IR), the black rectangle refers to the position of the floor sensor (FS). The dotted lines indicate the agent's view with the the three camera's sectors. (c) Agents starting positions within the nest.

distinguish the two types of environment. The nest is also indicated by a green light positioned at the centre of the nest. The foraging site is instead indicated by a red light positioned at a distance from the centre of the nest that varies from 100 cm to 110 cm. Both green and red light are positioned 6 cm above the floor and, when turned on, they are visible from everywhere within the arena.

The simulated robot is modelled as a circular object (of radius 3.5 cm like an e-puck robot) with left and right motors that can be independently driven forward or in reverse, allowing the simulated robot to turn in any direction. The agent maximum speed is 8 cm/s. The agent position is updated using the Differential Drive Kinematic equations described in (Dudek and Jenkin, 2000). The simulated robot is provided with eight infra-red sensors (IR^i with $i = \{0, .., 7\}$), which give the agent a noisy and non-linear indication of the proximity of an obstacle (in this task, an obstacle can only be another agent); a linear camera to see the lights; and a floor sensor (FS) positioned facing downward on the underside of the agent (see Figure 2b). The IR sensor values are extrapolated from look-up tables provided with the Evorobot* simulator (Nolfi and Gigliotta, 2010). The FS sensor can be conceived of as an IR sensor capable of detecting the intensity of grey of the floor. It returns 0 if the simulated robot is on white floor, 0.5 if it is on light grey floor, and 1 if it is on dark grey floor. The agents' camera has a receptive field of 30°, divided in three equal sectors, each of which has three binary sensors (C_i^B for blue, C_i^G for green, and C_i^R for red, with $i = \{1, 2, 3\}$ indicating the sector). Each sensor returns a value which is 0 if no light is detected, and 1 when a light is detected. The camera can detect coloured objects up to a distance of 150 cm. In this study, we assume that robots of a group share the same hardware structure. Note that supplementary materials concerning further details of the robot model, further graphs illustrating groups' performances, and videos of swarms of simulated robots operating in these task-allocation scenario can be found online at `http://users.aber.ac.uk/elt7/suppPagn/SI2015/suppMat.html`.

5 Controller

The agent controller is a continuous time-recurrent neural network (CTRNN) with 11 sensor neurons, 5 inter-neurons, and 2 motor neurons (Beer and Gallagher, 1992). The structure of the network is shown in Figure 2a. The values of sensory, internal, and motor neurons are updated using equations 1, 2, and 3:

$$y_i = gI_i;\ i = 1, \ldots, 11; \tag{1}$$

$$\tau_i \dot{y}_i = -y_i + \sum_{j=1}^{16} \omega_{ji}\sigma(y_j + \beta_j);\ i = 12, \ldots, 16; \tag{2}$$

$$y_i = \sum_{j=12}^{16} \omega_{ji}\sigma(y_j + \beta_j);\ i = 17, \ldots, 18. \tag{3}$$

In these equations, using terms derived from an analogy with real neurons, y_i represents the cell potential; τ_i the decay constant; g is a gain factor; I_i with $i = 1, \ldots, 11$ is the activation of the i^{th} sensor neuron (see Figure 2a for the correspondence between agent's sensors and sensor neuron); ω_{ji} the strength of the synaptic connection from neuron j to neuron i, β_j the bias term; $\sigma(y_j + \beta_j)$ the firing rate (hereafter, f_i), with $\sigma(x) = (1 + e^{-x})^{-1}$. All sensory neurons share the same bias (β^I), and the same holds for all motor neurons (β^O). τ_i and β_i with $i = 12, \ldots, 16$, β^I, β^O, all the network connection weights ω_{ij}, and g are genetically specified networks' parameters. At each time step, the firing rate of neurons 17 and 18 (i.e., f_{17} and f_{18}) are linearly scaled in [-1, 1] and then used to set the speed of the left and right wheels. Neurons' cell potentials are set to 0 when the network is initialised or reset, and equation 2 is integrated using the forward Euler method with an integration time step $\Delta T = 0.1$. Each network has 103 parameters (i.e., 90 connection weights, 5 decay constants, 7 bias terms, and 1 gain factor).

6 The Three Experimental Conditions

In this study, we investigated the problem of designing controllers for swarm of simulated robots engaged in the task described in Section 3 using three different experimental conditions. These conditions differ with respect to the nature of the evolutionary algorithm used to synthesise the agents' controller, the genetic relatedness of the group's members, and with respect to the evaluation function used to score the group performances. In the following, we first describe those aspects that are common to all the three experimental conditions and then we illustrate the distinctive characteristics of each of them.

During evolution, each group undergoes a set of $E = 4$ evaluation sequences (hereafter, *t-sequence*). A *t-sequence* is made of $V = 3$ trials. There are two different types of *t-sequence*: in *ABA-sequence* the agents experience *Env. A* in trial 1 and in trial 3, *Env. B* in trial 2; in *BAB-sequence* the agents experience *Env. B* in trial 1 and trial 3, *Env. A* in trial 2. Each group experiences twice each type of *t-sequence*. At the beginning of trial 1 of each *t-sequence*, the agents' controllers are reset, and each agent is randomly placed within an area corresponding to a sector of the nest. The nest is divided in 6 sectors. Each agent is randomly placed in one sector, and

randomly oriented in a way that the light can be within an angular distance of $\pm 36^{\circ}$ from its facing direction (see Figure 2c).

Each trial differs from the others in the initialisation of the random number generator, which influences the agents initial position and orientation, all the randomly defined features of the environment, and the noise added to motors and sensors. Within a trial, the group life-span is T=900 simulation cycles (with 1 simulation cycle lasting 0.1 s). Simulated robots are frozen (i.e., they do not move and do not contribute to the group fitness) if they exceed the arena limits (i.e., a circle of 120 cm radius, centred in the middle point between the nest and the foraging site). Trials are terminated earlier if all the agents are frozen, or the group exceeds the maximum number of collisions (i.e., 10). In trials following the first one of each *t-sequence* (trials 2, and 3), the agents are repositioned only if the previous trial has been terminated earlier, or with one or more agents frozen.

Each trial is divided into three phases. During *Phase 1*, which lasts 12 s, the green light is on and the red light is off. The simulated robots are required to stay within the nest. During *Phase 2*, which can last from a minimum of 47,5 s to a maximum of 52.5 s, the red light is on and the green light is off. During *Phase 2*, a group is required to behave according to the rules of the task. That is, in *Env. A*, the majority of agents has to visit the foraging site and the minority has to remain for the entire length of this phase in the nest. In *Env. B*, the majority of agents has to remain for the entire length of *Phase 2* in the nest and the minority has to visit the foraging site. A simulated robot is considered having visited the foraging site if, during *Phase 2*, it spends more than 100 consecutive time steps at less than 45 cm from the light indicating the foraging site. During *Phase 3*, which starts at the end of *Phase 2* and terminates at the end of the trial, the green light is on again and the red light is off. The agents that were foraging during *Phase 2* are required to return in the nest to rejoin their group mates. The agents can not see each other through the camera. Thus, any agent-agent interactions, including those that drive the task-allocation process, are based on the activations of the infra-red sensors. This is to keep our model as similar as possible to those that have investigated similar issues (see Quinn, 2001; Tuci and Trianni, 2012, 2014; Tuci, 2014).

6.1 First Experimental Condition (the *Evo-clonal* approach)

In the first experimental condition groups are homogeneous. That is, a group is formed using a single genotype from the evolving population of genotypes. Thus, each genotype generates "cloned" control software for all the simulated robots. An evolutionary algorithm using linear ranking is employed to set the parameters of the networks (Goldberg, 1989). Hereafter, the evolutionary approach of this experimental condition is called *evo-clonal* approach. We consider populations composed of $\gamma = 100$ groups, each composed of $\eta = 5$ individuals. The genotypes coding for the parameters of the agents' controllers are vectors comprising 103 real values chosen uniformly random from the range [0,1]. Each of the γ groups at generation 0 is formed by generating one random genotype and cloning it $\eta - 1$ times to obtain η identical genotypes. Generations following the first one are produced by a combination of selection with elitism, recombination, and mutation. For each new generation, the highest scoring genotype ("the elite") from the previous gen-

eration is retained unchanged, and used to form a new group. Each of the other $\gamma - 1$ new groups are generated by fitness-proportional selection from the 60 best genotypes of the old population. Each new genotype has a 0.3 probability of being created by combining the genetic material of two individuals of the old population. During recombination, one crossover point is selected. Mutation entails that a random Gaussian offset is applied to each real-valued vector component encoded in the genotype, with a probability of 0.04. The mean of the Gaussian is 0, and its standard deviation is 0.1. During evolution, all vector component values are constrained to remain within the range [0,1].

In this experimental condition, each solution is evaluated by a fitness function which rewards groups in which the agents remain within the nest during *Phase 1* and *Phase 3* and in which the tasks are allocated according to the rules of the game (i.e., majority of the agents on *task F* in *Env. A*, and majority of the agents on *task P* in *Env. B*). The average group evaluation score ($\bar{M}$) is computed as follows:

$$\bar{M} = \frac{1}{EV} \sum_{e=1;v=1}^{E,V} M_{ev}; \quad \bar{M} \in [0,7], E = 4, V = 3; \tag{4}$$

$$M_{ev} = \left[2 \times \left(\frac{\sum_{r=1}^{R} S_r^{ph1}}{R \times T^{ph1}} \times \frac{\sum_{r=1}^{R} S_r^{ph3}}{R \times T^{ph3}} \right) + U^{ph2} \right] \times P; \tag{5}$$

$$U^{ph2} = \begin{cases} 5 & \text{if agents are correctly allocated in } \textit{Phase 2}; \\ 2 \times \frac{N}{R} & \text{if agents are incorrectly allocated in } \textit{Phase 2}; \end{cases} \tag{6}$$

where $R = 5$ corresponds to the total number of agents in a swarm; S_r^{ph1} is the number of simulation cycles agent r spends within the nest during *Phase 1*; S_r^{ph3} is the number of simulation cycles agent r spends within the nest during *Phase 3*; T^{ph1} and T^{ph3} are the length (i.e., the number of simulation cycles) of *Phase 1* and *Phase 3*, respectively. The group collision penalty P is inversely proportional to the number of collisions, with $P = 1$ with no collisions, and $P = 0.4$ with 10 collisions in a trial. η corresponds to either the number of agents playing *task F*, during *Phase 2*, if the trial is in *Env. A*; or the number of agents playing *task P*, during *Phase 2*, if the trial is in *Env. B*.

6.2 Second Experimental Condition (the *Evo-aclonal-single* approach)

In the second experimental condition each group is formed of multiple genotypes (one for each group member) from the evolving population of genotypes. Each genotype generates the control software for only one agent. It follows that aclonal groups are heterogeneous because each group member has a controller derived from a different genotype. As in *evo-clonal* approach, an evolutionary algorithm using linear ranking is employed to set the parameters of the networks. Hereafter, the evolutionary approach of this experimental condition is called *evo-aclonal-single* approach. At generation 0, each of the $\gamma = 100$ groups is formed by generating $\eta \times \gamma$ random genotypes, with $\eta = 5$ agents in a group. As for the *evo-clonal* approach, the genotypes coding for the parameters of the agents' controller are vectors comprising 103 real values chosen uniformly random from the range [0,1]. For each new generation following the first one, the genotypes of the best group

("the elite") are retained unchanged and copied to the new population. Each of the genotypes of the other groups is formed by first selecting two old groups using roulette wheel selection from the 60 best groups of the old population. Then, two genotypes, each randomly selected among the members of the selected group are recombined with probability 0.3 to reproduce one new genotype. The resulting new genotype is mutated with probability 0.04. Mutation and recombination are applied in the same way as for the *evo-clonal* approach. This process is repeated to form $\gamma - 1$ new groups of η genotype each.

In this experimental condition, each solution is evaluated by a fitness function which scores the group performance using the metric $\bar{M}$ described in Section 6.1, and the metric L_r, referred to as *plasticity index*. L_r is a function of the frequency with which an agent executes each task in each environment, and it measures the agent's behavioural plasticity. The maximum L_r score can be achieved by an agent that executes both tasks in both types of environment with exactly the same frequency.

The average group evaluation score ($\bar{F}$) is computed in the following:

$$\bar{F} = \bar{M} + \bar{L}; \quad \bar{F} \in [0, 12]; \tag{7}$$

$$\bar{L} = \sum_{r=1}^{R} L_r;\ \bar{L} \in [0, R];\ L_r = \prod_{i=1}^{4} \Theta_r^i;\ L_r \in [0, 1]; \tag{8}$$

$$\Theta_r^i = \begin{cases} \frac{Q_r^i}{(0.25EV)} & \text{if } Q_r^i \leq (0.25EV); \\ 2 - \frac{Q_r^i}{(0.25EV)} & \text{otherwise;} \end{cases} \tag{9}$$

where Q_r^1 is the number of times the agent r performs *task F* in *Env. A*; Q_r^2 is the number of times the agent r performs *task P* in *Env. A*; Q_r^3 is the number of times the agent r performs *task F* in *Env. B*; and Q_r^4 is the number of times the agent r performs *task P* in *Env. B*.

Given the nature of the scenario, there are four possible task-environment combinations (i.e., *task F* in *Env. A*, *task P* in *Env. A*, *task F* in *Env. B*, and *task P* in *Env. B*). Within an agent's life-span, corresponding to EV trials, a fully generalist agent (i.e., an agent that has no preference for any of the four task-environment combinations) is expected to execute each combination $0.25EV$ times. Conversely, an agent that can only perform a single task regardless of the environment's type (i.e., a fully specialist agent), executes its preferred task $0.5EV$ times in each type of environment. Thus, 0.25 and 2 in equation 9 are normalisation factors that keep the Q_r^i values within the interval $[0, 1]$.

6.3 Third Experimental Condition (the *Evo-aclonal-multi* approach)

The third experimental condition is characterised by the use of the evolutionary multi-objective optimisation algorithm NSGA-II (see Deb et al., 2002, for a detailed description of the algorithm). Hereafter, the evolutionary approach of this experimental condition is called *evo-aclonal-multi* approach. As in *evo-aclonal-single* approach, groups are heterogeneous. That is, each group is formed from multiple genotypes (one for each group member) from the evolving population of genotypes. Each genotype generates the control software for only one agent. At generation

zero, a random population of $\gamma = 60$ heterogeneous groups is created. Each group is made of $\eta = 5$ different genotypes coding for the parameters of the agents' controller as explained in Section 6.1. During evaluation, each group is scored against two objectives: $\bar{M}$ (see Equation 4) and $\bar{L}$ (see Equation 8). After evaluation, each group is ranked according to a criteria that compares groups with respect to their scores in both objectives (see Deb et al., 2002, for details). Following generations are created using the algorithm illustrated in (Deb et al., 2002), using binary tournament selection, probability of mutation 0.04, and probability of recombination 0.3. Mutation and recombination are applied in the same way as for the *evo-clonal* approach. Minor modifications are introduced with respect to the algorithm described in (Deb et al., 2002) in order to introduce group selection. In our implementation of NSGA-II, genotypes belonging to the same group have equal fitness. Thus, when we create a new population of size $\gamma \times \eta$ genotypes, we first select the group based on its rank, and within a group we randomly choose one genotype, as for the *evo-aclonal-single* approach. During the groups' selection process, groups with equal rank are selected using the *crowding distance* (i.e., the Euclidean distance between groups in the two-dimensional fitness space). The higher the average crowding distance of a group, the higher its probability to be selected.

7 Results

We ran 20 differently seeded evolutionary simulations for each experimental condition. Each run lasted 2500 generations. The objective of this study is to verify, first, that the two evolutionary approaches that work with heterogeneous groups (i.e., the *evo-clonal* approach and the *evo-aclonal-multi* approach) outperform the evolutionary approach working with homogeneous groups (i.e., the *evo-clonal* approach). Second, we intend to investigate whether the *evo-aclonal-multi* approach and the *evo-aclonal-single* approach can generate behaviourally plastic agents, and eventually if any of these two aclonal approaches can be preferred to the clonal one to design controllers for a swarm of simulated robots engaged in task-allocation scenarios.

To verify our hypothesis we ran several post-evaluation tests[3]. The first test aims to estimate, in a broader range of initial conditions, the effectiveness of group strategies generated by artificial evolution. At the end of the evolutionary phase, we selected for post-evaluation the most promising groups for each experimental condition. For the *evo-clonal* approach and the *evo-aclonal-single* approach we re-evaluated the best 5 groups of each of the last 1000 generations. For the *evo-aclonal-multi* approach we re-evaluated, for each of the last 1000 generations, all the Pareto-optimal (i.e., non-dominated) groups. In these tests, the metric $\bar{M}$ is used to quantify the extent to which a group complies with the rules of the game (see Equation 4 for details).

[3] In all post-evaluation tests described in this Section, each single group undergoes a set of E=80 differently seeded *t-sequences* (40 *ABA-sequence*, and 40 *BAB-sequence*), each made of V=3 trials, for a total of 240 trials, 120 trials in *Env. A* and 120 trials in *Env. B*. Each *t-sequence* differs from the others in the initialisation of the random number generator, which influences the agents initial position and orientation at trial 1 and during repositioning, all the randomly defined features of the environment, and the noise added to motors and sensors.

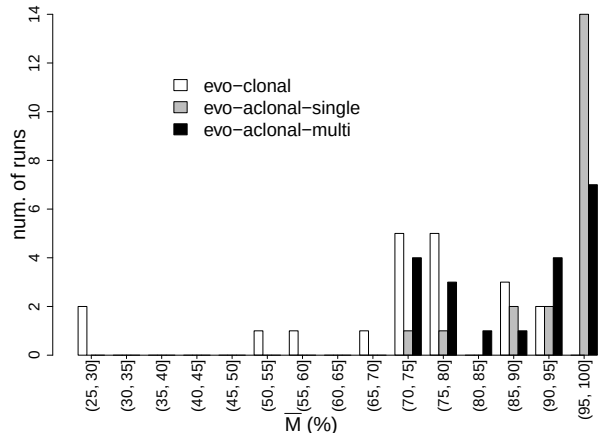

Fig. 3 Graph showing the performance ($\bar{M}$) of the best evolved group of each evolutionary run of each experimental condition. White bars refer to the *evo-clonal* approach, grey bars refer to the *evo-aclonal-single* approach, and black bars refer to the *evo-aclonal-multi* approach. Each bar shows the number of runs per condition whose best performing group falls into the interval indicated on the x-axis. Values on the x-axis represent percentages of the optimal evaluation score $\bar{M} = 7$.

The graph in Figure 3 illustrates the average post-evaluation performance of the best group for each evolutionary run and experimental condition. The bars refer to the number of runs per condition whose best performing group falls into the interval indicated on the x-axis. Values on the x-axis represent percentages of the optimal evaluation score $\bar{M} = 7$. The results of this test clearly indicate that the evolutionary approaches that work with heterogeneous groups (see Figure 3 grey and black bars) outperform the *evo-clonal* approach working with homogeneous groups (see Figure 3 white bars). Table 1 shows mean, standard deviation, and median of the performance $\bar{M}$ computed on the 20 best groups of each experimental condition. Each measure in the Table shows that both the *evo-aclonal-single* approach and the *evo-aclonal-multi* approach outperform the *evo-clonal* approach. Moreover, the difference between *evo-aclonal-single* approach and *evo-clonal* approach as well as the differences between *evo-aclonal-multi* approach and *evo-clonal* approach are statistically significant (Wilcoxon rank sum test, test comparing *evo-clonal* approach and *evo-aclonal-single* approach $p < 0.001$; test comparing *evo-clonal* approach and *evo-aclonal-multi* approach $p < 0.01$). From a statistical point of view, there is enough evidence to prefer the experimental conditions working with heterogeneous groups over the condition working with homogeneous groups.

Table 1 Table showing median, mean and standard deviation of performance ($\bar{M}$) of the best evolved groups for each run of each experimental condition.

	median	mean	s.d.
evo-clonal approach	5.26	5.02	1.21
evo-aclonal-single approach	6.81	6.58	0.50
evo-aclonal-multi approach	6.45	6.12	0.72

7.1 On the Behavioural Plasticity of Agents of Heterogeneous Groups

The results of the first post-evaluations have confirmed that, as in (Quinn, 2001; Tuci and Trianni, 2014; Tuci, 2014), the aclonal approaches outperform the clonal approach. We have shown that evolving homogeneous groups is a less efficient way to generate solutions for swarms of agents engaged in this task-allocation scenario than evolving heterogeneous groups. However, as stated in Section 2, our primary objective is to investigate whether the aclonal approaches can generate successful groups made of behaviourally plastic agents.

As mentioned in Section 1, in this study the agents' behavioural plasticity is evaluated in circumstances in which the agents are required to adjust their task preferences in response to the behaviour of their group mates. During evolution, we have measured behavioural plasticity with the metric $\bar{L}$, corresponding to the frequency with which an agent performs each task in each environmental condition (see Section 6.2 and equation 8). We are aware that the metric $\bar{L}$ has some limitations, as it can not detect "false negative" (i.e., agents which are behaviourally plastic but that do not act as behaviourally plastic agents due to group constraints). In a heterogeneous group, it could be that, due to the characteristics of the group members, a potentially plastic and generalist agent may not exhibit its plasticity because, for the benefit of the group, it ends up repeatedly performing those tasks that specialist group mates are not capable of doing. During the design phase, we could not interfere with the group assemblage process, which is entirely delegated to the evolutionary machinery. Moreover, adding further evaluations to detect false negatives would have increased the computational time required to generate successful solutions. Thus, during evolution, $\bar{L}$ seemed a reasonable compromise between the possibility to fail to adequately reward heterogeneous groups with potentially plastic agents, and the necessity to limit the computational time required to run an evolutionary process.

During post-evaluation, we take advantage of the possibility to interfere with the group assemblage process in order to develop metrics of an agent's behavioural plasticity that are not subject to the effect of false negative. In particular, we require an agent to operate in a homogeneous group, where its behavioural plasticity can be reliably estimated using the group performance $\bar{M}$. Recall that, if under the conditions of this task, a homogeneous group is successful, then its agents are behaviourally plastic. This follows from the fact that, in spite of being clones, the agents efficiently employ task-allocation mechanisms to distribute the resources to *task P* and *task F* according to the rules of the task explained in Section 3. Thus, for this set of post-evaluations, we chose to measure the behavioural plasticity of agents generated aclonally through observing the performance of the homogeneous groups generated by cloning each of them. Specifically, we post-evaluated in a homogeneous group (i.e., in a group of clones) each agent originally part of heterogeneous groups selected from all runs of both *evo-aclonal-single* approach and *evo-aclonal-multi* approach that at the first post-evaluation test showed a group performance $\bar{M} > 75\%$. The performance of homogeneous groups (measured using the metric $\bar{M}$) is considered a quantitative estimate of the behavioural plasticity of the agent generating the group. The higher the $\bar{M}$ score, the more behaviourally plastic the agent which generated the group.

The graph in Figure 4 shows the performance of the best 2000 homogeneous groups generated by cloning agents originally forming: i) heterogeneous groups

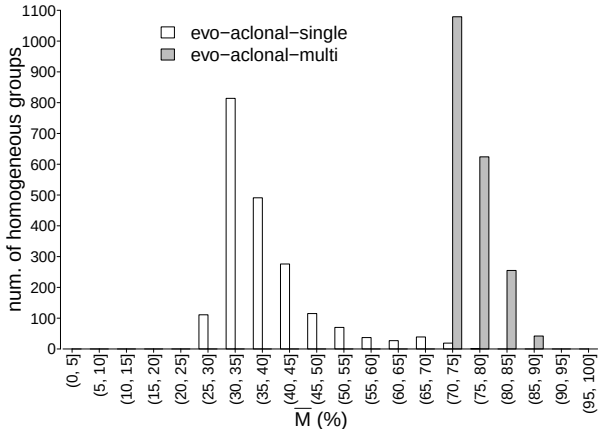

Fig. 4 Graph showing the performances ($\bar{M}$) of the best 2000 homogeneous groups generated by cloning agents originally forming: i) heterogeneous groups generated by the *evo-aclonal-single* approach (white bars); ii) heterogeneous groups generated by the *evo-aclonal-multi* approach (grey bars). Each bar shows the number of homogeneous groups whose performance falls into the intervals indicated on the x-axis. Values on the x-axis represent percentage of the optimal evaluation score $\bar{M} = 7$.

generated by the *evo-aclonal-single* approach (white bars); ii) heterogeneous groups generated by the *evo-aclonal-multi* approach (grey bars). Each bar refers to the number of homogeneous groups whose performance falls into the intervals indicated on the x-axis. The graph clearly shows that, when re-evaluated in homogeneous groups, the agents generated by the *evo-aclonal-multi* approach perform better than the agents generated by the *evo-aclonal-single* approach. Since all the 2000 selected agents generated by the *evo-aclonal-multi* approach managed to form homogeneous groups with a performance higher than 70% of the optimal $\bar{M}$ score, we conclude that these 2000 agents demonstrated to be sufficiently plastic to be able to adjust their behaviour in response to the preference of their group-mates. In this particular test, these agents showed that their behavioural plasticity can overcome the initial symmetry due to the agents being clones in order to trigger an effective task allocation that complies with the rule of this scenario.

Figure 4 shows that agents generated by the *evo-aclonal-single* approach perform very poorly when re-evaluated in homogeneous groups (see Figure 4, white bars). Only 193 agents out of 2000 managed to get a score higher than 50% of the optimum score, and only 20 agents generated homogeneous groups that scored more than 70% of the optimal $\bar{M}$ value. We also notice that these 20 agents, which demonstrated to be behaviourally plastic by generating sufficiently well performing homogeneous groups, all come from a single evolutionary run.

Given the results of this test, we conclude that the *evo-aclonal-multi* approach is more effective than the *evo-aclonal-single* approach in generating heterogeneous groups that are effective in this task-allocation scenario, and in generating agents that proved to be highly behaviourally plastic. In the nest Section, we will try to account for the differences between the *evo-aclonal-single* approach and the *evo-aclonal-multi* approach in generating behaviourally plastic agents, by exploring the effects of the metric $\bar{L}$ on the evolutionary processes of both aclonal conditions.

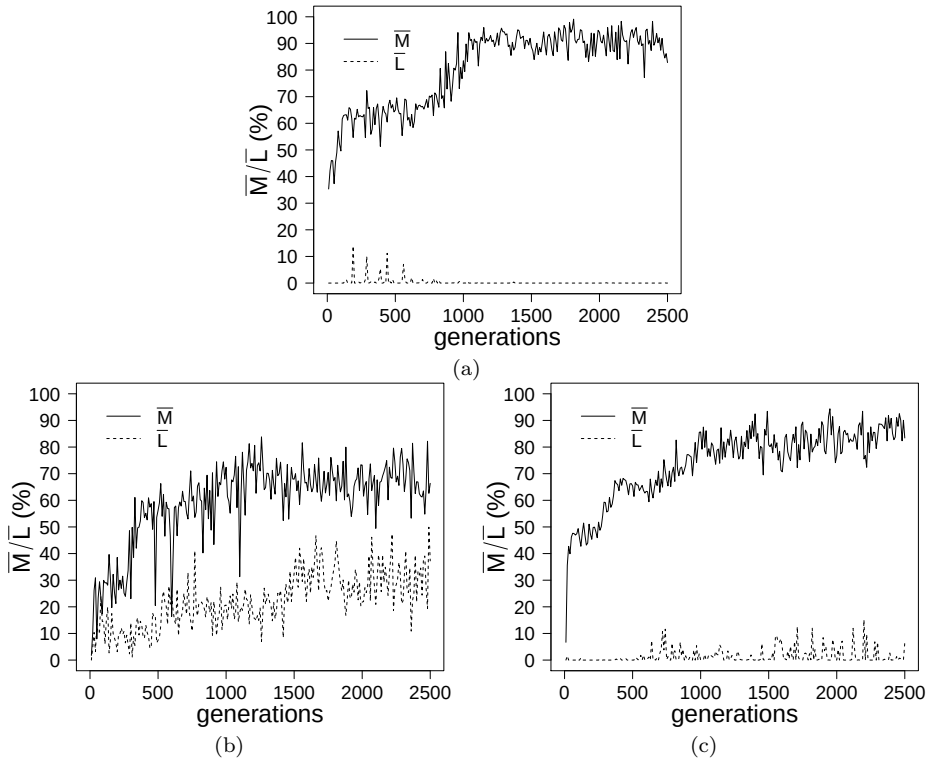

Fig. 5 Graphs showing the group performance $\bar{M}$ (continuous lines) and the group behavioural plasticity $\bar{L}$ (dashed lines) for each best group from generation 0 to generation 2500. Graph in a) refers to the best groups of one of the most successful evolutionary run of the *evo-aclonal-single* approach. Graphs b) and c) refer to one of the most successful evolutionary run of the *evo-aclonal-multi* approach from which we selected for re-evaluation the non-dominated groups with the highest performance at metric $\bar{L}$ (graph b) and non-dominated groups with the highest performance at metric $\bar{M}$ (graph c).

7.2 Initial Qualitative Analysis of the Evolutionary Dynamics of the *Evo-aclonal-single* approach and the *Evo-aclonal-multi* approach

In this section, we take a first step towards the exploration of the evolutionary dynamics of the aclonal approaches. We acknowledge that it is relatively difficult to analyse, compare, and interpret the evolutionary dynamics of an ESR experiment. For example, the complex and dynamic nature of neural network controllers makes it hard to trace behavioural traits back to their genes. Moreover, the use of the cross-over operator makes it difficult to reconstruct the evolutionary history of specific genes. Due to the complexity of this investigation, in this study, we limit our observations to simple qualitative data that nevertheless point to interesting phenomena. In particular, we provide further evidence indicating that the metric $\bar{L}$, used both in the *evo-aclonal-single* approach and *evo-aclonal-multi* approach to generate selective pressures to favour swarms of behaviourally plastic over swarms of specialist agents, only in the *evo-aclonal-multi* approach manages to steer evolution towards the desired outcomes. This evidence comes from the results of a

series of tests in which we re-evaluated all the best groups of each generation of those evolutionary runs of the *evo-aclonal-single* approach and of the *evo-aclonal-multi* approach that have been successful in generating heterogeneous groups with a group performance higher than 75% of the optimum score, and in generating behaviourally plastic agents whose homogeneous group performance score is higher than 50% of the optimum score.

During these tests, we have measured the group performance using the metric $\bar{M}$ and the metric $\bar{L}$. The graph in Figure 5a shows the evolutionary trajectory of one particular successful run of the *evo-aclonal-single* approach. We can notice that the group plasticity drops to zero from about generation 900 till the end of the evolution (see Figure 5a dashed line). In other words, the highest fitness groups of the most successful runs of the *evo-aclonal-single* approach are made of agents that act as specialists. At this point, we can not exclude the presence of false negatives (i.e., agents that are plastic but that repress their plasticity for the benefit of the group). However, even if the dashed-line curve shown in Figure 5a would hide groups with agents that repress their plasticity, the emergence of this repressed plasticity can not be accounted for by the fitness component $\bar{L}$, since as mentioned above, this fitness component can not select for a characteristic that does not express itself through the agents' behaviour. We conclude that the evolutionary trend shown in Figure 5a, and repeatedly observed in almost all runs of the *evo-aclonal-single* approach, is a clear sign of the very limited influence that the fitness component $\bar{L}$ exerts on the dynamics of *evo-aclonal-single* approach evolutionary processes. As a result of this, the fitness component $\bar{L}$ generally fails to steer evolution toward the emergence of groups with plastic agents.

Contrary to the *evo-aclonal-single* approach, in *evo-aclonal-multi* approach the multi-objective approach generates different evolutionary dynamics that do not penalise behaviourally plastic agents. Graphs in Figure 5b and 5c show the evolutionary trajectories of a successful evolutionary run of the *evo-aclonal-multi* approach. Figure 5b refers to the evolutionary trajectories generated by selecting the non-dominated group with the highest performance at metric $\bar{L}$. Figure 5c refers to the evolutionary trajectories generated by selecting the non-dominated group with the highest performance at metric $\bar{M}$. Both graphs shows that a certain amount of behavioural plasticity is represented in both best groups that, at each generation, are found at the opposite ends of the Pareto front (see Figure 5b and 5c dashed lines). Other successful runs show similar evolutionary trends. This indicates that, in the *evo-aclonal-multi* approach, the multi-objective approach generates evolutionary dynamics in which the genes of those agents, acting as behaviourally plastic, are kept into the gene pool of the evolving populations.

In conclusion, the results of the tests illustrated in this Section indicate that when the metric $\bar{L}$ operates as a distinctive objective in an evolutionary multi-objective optimisation approach rather then as an additive component of a single objective evolutionary approach, it clearly affects the evolutionary dynamics by steering evolution toward the emergence of groups with behaviourally plastic agents.

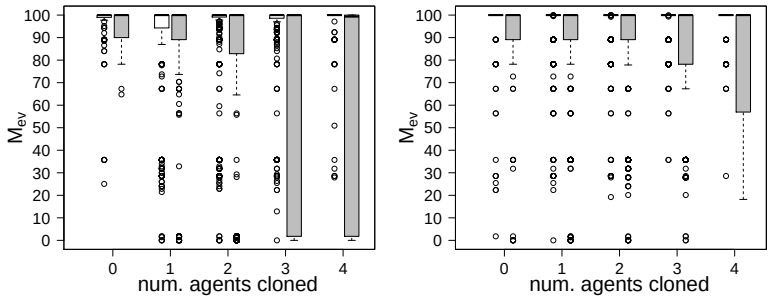

Fig. 6 Graphs showing the results of the replacement test for a) a group generated by the *evo-aclonal-single* approach; b) a group generated by the *evo-aclonal-multi* approach. In both graphs, white boxes refer to the performances in *Env. A*, while grey boxes refer to the performances in *Env. B*. On the x-axis, labels refer to the number of agents replaced by the most plastic individual of the group. Each point in the boxes refers to the group fitness in a single trial. The group fitness is measured through the percentage of maximum score $M_{ev} = 7$ achievable by a group in a single trial. Boxes represent the inter-quartile range of the data, while dashed horizontal bars inside the boxes mark the median values. The whiskers extend to the most extreme data points within 1.5 times the inter-quartile range from the box. Empty circles mark the outliers.

7.3 Further evidence of the behavioural plasticity of agents generated by the *evo-aclonal-multi* approach

We conclude the results section with a series of post-evaluation tests, which provide further quantitative evaluation of the behavioural plasticity of agents generated with the aclonal approaches. The aim of these tests is to illustrate that plastic agents generated aclonally are, to some extent, capable of operating in a variety of conditions, which differ in term of characteristics of group-mates as well as for the cardinality of the group.

We chose, for each experimental condition, a very successful group ($\bar{M} > 80\%$ when operating in its heterogeneous state) with at least one very plastic individual. Then, we ran two different types of test: the replacement test, in which the most behaviourally plastic agent of the group is required to progressively replace all group mates; and the scalability test, in which the most behaviourally plastic agent of the group is used to generate new clones that increase the cardinality of the group, up to a group of eleven agents.

The replacement test

In the replacement test, we progressively select one, two, and three agents to be replaced by the most plastic individual of the group (hereafter, *target agent*). We systematically considered all the possible combinations of the four agents selected for substitution one/two/three at a time without repetition. The performance of the "hybrid" group (i.e., a group in which some agents are clones and others are not) is recorded using the metric ($\bar{M}$).

The results of the replacement test are shown in Figure 6. Note that, boxes in correspondence of x-tick 0 refer to the condition in which the group operates in its normal heterogeneous set-up. Boxes in correspondence of x-tick 4 refer to the condition in which the group operates in an homogeneous set-up (i.e., all agents

are clones of the most plastic agent of the group). Figure 6a refers to a group generated by the *evo-aclonal-single* approach; Figure 6b refers to a group generated by the *evo-aclonal-multi* approach. The graphs in Figure 6 clearly indicate that, first, for the majority of the test conditions, the performance of both groups tend to be relatively good. Second, the group generated by the *evo-aclonal-multi* approach performs better than the group generated by the *evo-aclonal-single* approach. For example, for the group generated by the *evo-aclonal-multi* approach the performance drop is more pronounced only in correspondence of the fully homogeneous state when the group operates in *Env. B* (see Figure 6b, grey boxes, for x-tick 4). For the group generated by the *evo-aclonal-single* approach, the *Env. B* proved to be particularly challenging for the last two conditions (see Figure 6a, grey boxes, for x-tick 3 and 4). Third, for both groups the performance tends to moderately decrease moving from the fully heterogeneous state (boxes in correspondence of the x-tick 0) to the fully homogeneous state (boxes in correspondence of the x-tick 4). Fourth, both groups tend to perform better in one environment than the other. These four elements represent general characteristics that, with minor variations, have been observed in various other aclonally generated groups with similar characteristics in terms of performance and of agents plasticity. Thus, they are representative of what most successful heterogeneous groups, with plastic agents, can do when facing similar circumstances as those considered by the replacement test.

The scalability test

In the scalability test, we progressively added one extra agent to the group by cloning the most plastic agent of the group. The performance of each group is recorded using the metric ($\bar{M}$). We considered groups of up to 11 agents. We decided not to proceed further as this would require a significant restructuring of the scenario, including changes to the diameter of the nest and foraging site to fit all the agents, an increase of the maximum time for completion of each phase of the task, and changes to other elements of the original set-up. Variations applied to the spatio-temporal relationship of the task scenario, like those mentioned above, tend to have a disruptive effect on group performance, as agents operate according to temporal dynamics evolved to cope with the events detailed in Section 4 and Section 6. The results shown in Figure 7 are nevertheless informative of the capability of plastic agents to operate in larger groups.

The graph in Figure 7a refers to the performance of the most successful group generated by the *evo-clonal* approach. In Figure 7a, all boxes refer to the performance of homogeneous groups. The performance of this group can be used as a comparison to evaluate the effectiveness of plastic agents generated by aclonal conditions in order to cope with scalability issues. Figure 7b, and Figure 7c refer to the performance of a successful group generated by the *evo-aclonal-single* approach and *evo-aclonal-multi* approach, respectively. For the *evo-aclonal-single* approach and *evo-aclonal-multi* approach, the groups tested were also used for the replacement test. In Figure 7b and Figure 7c, the boxes in correspondence of x-tick 5 refer to the condition in which the group operates in its normal heterogeneous state with 5 agents. The boxes in correspondence of x-tick bigger than 5 refer to the performances of hybrid groups (i.e., some agents are clones, others are not clones), in which the extra agents are clones of the most plastic agent of the group.

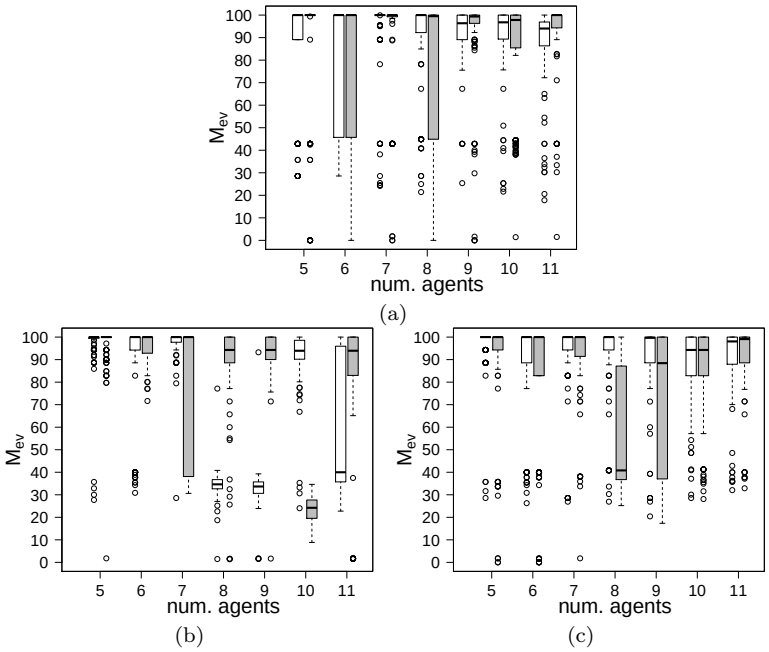

Fig. 7 Graphs showing the results of the scalability test for a) a group generated by the *evo-clonal* approach; b) a group generated by the *evo-aclonal-single* approach; c) a group generated by the *evo-aclonal-multi* approach. In all graphs, white boxes refer to the performances in *Env. A*, while grey boxes refer to the performances in *Env. B*. Label on the x-axis refer to the number of agents in the groups. Each point in the boxes refers to the group fitness in a single trial. The group fitness is measured through the percentage of maximum score $M_{ev} = 7$ achievable by a group in a single trial. Boxes represent the inter-quartile range of the data, while dashed horizontal bars inside the boxes mark the median values. The whiskers extend to the most extreme data points within 1.5 times the inter-quartile range from the box. Empty circles mark the outliers.

The graphs in Figure 7 clearly indicate that, first, for the majority of the test conditions, the performance tend to be relatively good. Second, both aclonally generated groups seem to handle scalability in a rather similar way to the groups generated by the *evo-clonal* approach. For all three groups, performances at the scalability test do not progressively decrease with the increment of the group size. There are instead specific conditions, different for each group, in which the groups find more difficult to successfully complete the task. For both aclonally generated groups, these conditions are limited to a single environment (e.g., Figure 7b, grey box, 7-agent and 11-agent groups, and Figure 7c, grey box, 8-agent and 9-agent groups). Visual inspection of the behaviour of these groups showed that, in those conditions where the performance drop, the groups seem to be highly sensitive to the agents' initial relative positions, which very often prevent the group from developing those virtuous dynamics underpinning the task-allocation process.

As for the replacement test, the patterns observed in Figure 7 relative to the scalability test represent general characteristics that, with minor variations, have been observed in various other groups with similar characteristics in terms of performance and agents' plasticity. Thus, they are representative of what most

successful homogeneous and heterogeneous groups with plastic agents can do when facing similar circumstances of those considered by the scalability test.

8 Conclusions

In 2001, a short conference paper (Quinn, 2001) proposed a rather counter-intuitive hypothesis suggesting that, in an ESR task-allocation scenario, it is more effective to obtain successful homogeneous swarms by evolving heterogeneous swarms (i.e., by using an aclonal approach) rather than through the evolution of homogeneous swarms (i.e., by using a clonal approach). Since in homogeneous swarms, the allocation of tasks is purely based on a dynamic negotiation between generalist agents, this hypothesis can be interpreted as claiming that the evolution of heterogeneous swarms represents a better approach to the evolutionary design of swarms in which the agents are generalists and capable of adapting their task preferences to the preferences of their group mates. This hypothesis, based on the beneficial effects of specialisation, was originally tested in a very simple two-agent coordination task, and the validation of its underlying principles was left to further investigations.

A recent study turned its attention toward this hypothesis with data that confirmed the advantages of aclonal over clonal approaches for the evolutionary design of successful robotic swarms engaged in task-allocation scenarios, but casting serious doubts on the potentialities of aclonal approaches in generating generalist agents (Tuci, 2014). Building upon the evidence in (Tuci, 2014), we showed that, by using an evolutionary multi-objective rather than a single objective optimisation algorithm, aclonal approaches facilitate the design of controllers for a swarm of simulated robots operating in a relatively complex task-allocation scenario and, at the same time, allow the designer to generate behaviourally plastic agents that can perform different tasks according to the swarm needs. We showed that the *evo-aclonal-multi* approach should be preferred to the *evo-aclonal-single* approach as with the former approach we can more easily generate behaviourally plastic agents (see Section 7.1). As suggested in (Quinn, 2001), we also found that specialisation in aclonal approaches represents a strong catalyst that accelerates the emergence of successful swarms. However, only in combination with an evolutionary multi-objective optimisation algorithm the aclonal approach seems to be able to systematically retain behaviourally plastic agents into the evolving populations, generating swarms in which specialist and generalist agents successfully cooperate to solve the group task (see Section 7.2 for details). We also showed that behaviourally plastic agents generated by the *evo-aclonal-multi* approach can successfully operate in hybrid groups in which their clones progressively replace all the original group mates, as well as in groups in which their clones are used to increase the group size up to more than twice the original group size (see Section 7.3).

The parameters of this model (e.g., the number of hidden nodes of the artificial neural network, or the mutation rate for the evolutionary algorithm, etc.) have been set to arbitrary values that, to the best of our knowledge, make the system potentially capable of generating the desired solutions without creating undesirable overheads (e.g., an increment of the computational time required to run an evolutionary process). During early stages of this research work, we have experimentally explored a relatively small set of the potentially infinite parameters space, observing no relevant differences on the results of the comparative analysis

discussed in this study. We believe that the conclusions we drew from the results of our research work can be generalised to a broader set of parameters, as they mainly pertain to the *modus operandi* of the clonal and aclonal approaches rather than to their specific implementation. The rationale behind the choice of using five-robot groups is the following: a five-robot group is the smallest, and the computationally less expensive, group in which the "minority" can be represented by more than one robot (i.e., by a two-robot group). By providing evolution with the means to represent the concept of "minority" by relaying on a two-robot group, we intended to facilitate the evolution of mechanisms that could potentially work for representation of the concept of "minority" with more than two robots, in case of larger groups. This is an issue that bears upon the variability and scalability of successful groups' strategies.

We believe the main contribution of this study is in showing that, within this task-allocation scenario, the aclonal approach in combination with a multi-objective optimisation algorithm generates a sufficiently large number of behaviourally plastic agents without having to introduce elements that may negatively impact on the computational time required to generate successful solutions. There are a variety of possible options to be considered in order to steer the evolutionary trajectories of single-objective aclonal runs towards the emergence of successful robotics swarms made of generalist agents. Many of them seem to come with undesirable additional computational overhead. For example, selective pressures in favour of generalist agents could be introduced by evaluating each single solution in multiple different groups. However, this option requires an increase in the number of evaluation trials with a consequent increase of the computational time required to evaluate single solutions. Alternatively, the group task could be designed in such a way that optimal performances can be achieved only by swarms with at least some generalist agents. This can be obtained either by introducing a larger set of operating conditions or by contemplating circumstances like fault agents whose role needs to be replaced by other agents. Both options, to be effective, tend to introduces extra computational time. The reader may have noticed that our aclonal multi-objective approach avoided these computational overheads by evaluating each solution in a single group, and by contemplating a scenario that does not require the presence of fully behaviourally plastic agents for a swarm to reach the optimal performance. This means that a swarm can be successful even if it is made of specialist agents with at least one agent that can perform multiple tasks but not in all environmental conditions (see Section 3 for more details). In spite of this, we demonstrated that generalist agents (i.e., agents that can do both tasks in both types of environment) emerged thanks to the evolutionary dynamics generated by the multi-objective optimisation algorithm.

In this study, the agents behavioural plasticity has been tested only in circumstances where the re-distribution of resources was dictated by variability within the composition of the swarm. We hope the results illustrated in this study will represent a valuable contribution to the research community interested in ESR. We indicated an alternative way that can be potentially helpful in designing, using evolutionary techniques, swarm robotics systems and behaviourally plastic agents capable of operating in more challenging environmental conditions in which the re-distribution of resources through a task-allocation process are dictated by multiple sources of environmental variability.

In recent years, the literature on task-allocation in swarm robotics has significantly grown. Some research works focus on scenarios in which a task can be either completely executed by generalist robots or partitioned in a such a way that different robots execute different parts of the task (Labella et al., 2006; Pini et al., 2011, 2013b,a; Ferrante et al., 2015). These studies primarily investigate the environmental conditions under which task partition is beneficial to the swarm in spite of the costs related to the coordination of specialist agents (see Pini et al., 2013b). The contribution of these studies is in illustrating the effectiveness of distributed mechanisms in generating complex group level responses, such as the division of a task into sub-tasks that can be tackled separately. Moreover, in some studies individual mechanisms allow single agents to act either as generalist or as specialist in response to the environmental conditions in which the group is required to operate (e.g., see Ferrante et al., 2015). Other studies investigate scenarios in which task allocation is assumed to be beneficial, either because the experimental scenario features concurrent tasks (see Ducatelle et al., 2009; Brutschy et al., 2012), as in this study, or because a group task is organised in sub-tasks that have to be executed in a predefined order (see Brutschy et al., 2014). In these studies, the focus is generally on how the competencies and characteristics of the single agents bear upon the group performance. For example, Brutschy et al. (2012) look at how the spatial and temporal distribution of tasks bears upon the agents' capabilities to improve their behavioural skills in executing specific tasks. The above mentioned body of literature is based on the use of finite state machine type of controllers in which stochastic processes regulate individual task preferences according to the agent history of interaction with the environment. In spite of sharing a common interest on the issue of task allocation for swarm robotics systems, the use of different design methods makes this literature only marginally relevant to this study. Thus, in the remaining of this section, we mainly review relevant works that employ evolutionary computation techniques and artificial neural networks as design methods.

In the ESR literature, only a few research works target task-allocation scenarios, and the majority of them look at tasks in which robots have to differentiate in leaders and followers. Some of these studies are interested in the means to achieve a more efficient allocation of roles rather than in the effects of the genetic composition of the teams (Gigliotta et al., 2014). Other research works, focused on the relationship between the genetic composition of the team and the task requirements, indicate in which operating conditions heterogeneous swarms appear to be more efficient than homogeneous swarms and vice versa (see, e.g., Ijspeert et al., 2001; Waibel et al., 2009). This interesting body of literature is nevertheless only partially relevant to this work. Therefore, we refer the reader to (Nitschke et al., 2007) for an extensive review of these studies. More relevant for this study is another series of research works, which do not explicitly target swarm robotics and task allocation. Instead, these works focus on the development of methods that facilitate the emergence of behavioural diversity into the population of evolving agents through the use of fitness functions that score agents for the novelty of their behaviour rather than for how effective they are in accomplishing specific goals (Stanley and Miikkulainen, 2004). Behavioural diversity is demonstrated to reduce the undesired effects of premature convergence in evolutionary computation and represents an alternative for the design of open-ended evolutionary processes (Lehman and Stanley, 2008). This literature, which is extensively and

nicely reviewed in (Doncieux and Mouret, 2014), shared an interest with this study in improving existing design methods and generating new alternative solutions to facilitate the task of engineering self-organising systems. Particularly relevant to us is the study described in (Mouret, 2011) which, like this study, emphasises the advantages of using multi-objective optimisation algorithms to exploit the benefits of rewarding agents both for their behavioural diversity and for specific task-dependent objectives. The results shown in (Mouret, 2011) suggest possible extensions of our work. The concept of behavioural diversity could be exploited to integrate task-independent as well as task-dependent objectives in order to improve the efficiency of aclonal multi-objective optimisation algorithms to generate successful swarms and behaviourally plastic agents for complex task-allocation scenarios.

We began this study by acknowledging that scenarios requiring complex spatio-temporal organisation and task allocation in swarm robotics systems need to be addressed by generating both effective mechanisms for the distribution of robots to tasks and behaviourally plastic agents that can perform different tasks for the benefit of the swarm. Recent studies have emphasised the need to engineer self-organisation through investigations focused on the effects of design choices offered by the evolutionary approach on the quality of the solutions (Dancieux et al., 2015; Trianni, 2014). In line with this perspective, this study contributes to the development of a principled methodological approach to the design of effective swarm robotics systems operating in a task-allocation scenario. We showed that successful group strategies and individually plastic agents can be generated using an aclonal evolutionary multi-objective optimisation algorithm. We believe that the evidence in this study will facilitate the engineering of more efficient swarm robotics systems capable of operating in complex scenarios requiring task allocation.

References

M. Allwright, N. Bhalla, H. El-faham, A. Antoun, C. Pinciroli, and M. Dorigo. SRoCS: Leveraging stigmergy on a multi-robot construction platform for unknown environments. In M. Dorigo, M. Birattari, S. Garnier, H. Hamann, M. Montes de Oca, C. Solnon, and T. Stützle, editors, *Proceedings of the* 9^{th} *International Conference on Swarm Intelligence*, volume 8667 of *LNCS*, pages 158–169. Springer, 2014.

R. D. Beer and J. C. Gallagher. Evolving dynamic neural networks for adaptive behavior. *Adaptive Behavior*, 1(1):91–122, 1992.

A. Brutschy, N.-L. Tran, N. Baiboun, M. Frison, G. Pini, A. Roli, M. Dorigo, and M. Birattari. Costs and benefits of behavioral specialization. *Robotics and Autonomous Systems*, 60(11):1408–1420, 2012.

A. Brutschy, G. Pini, C. Pinciroli, M. Birattari, and M. Dorigo. Self-organized task allocation to sequentially interdependent tasks in swarm robotics. *Autonomous Agents and Multi-Agent Systems*, 28(1):101–125, 2014.

S. Dancieux, N. Bredeche, J-B. Mouret, and A.E. Eiben. Evolutionary robotics: what, why, where to. *Frontiers in Robotics and AI*, 2:1–18, 2015.

K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2), 2002.

S. Doncieux and J-B. Mouret. Beyond black-box optimization: a review of selective pressures for evolutionary robotics. *Evolutionary Intelligence*, 7(2):71–93, 2014.

M. Dorigo and E. Şahin. Guest editorial. Special issue: Swarm robotics. *Aut. Rob.*, 17(2–3):111–113, 2004.

M. Dorigo, E. Tuci, R. Gross, V. Trianni, H.T. Labella, S Nouyan, J-L. Deneubourg, G. Baldassarre, S. Nolfi, F. Mondada, D. Floreano, and L.M. Gambardella. The swarm-bots project. In E. Şahin and W.M. Spears, editors, *Proceedings of the* 1^{st} *International Workshop on Swarm Robotics*, volume 3342 of *LNCS*, pages 31–44. Springer, 2004.

M. Dorigo, M. Birattari, and M. Brambilla. Swarm robotics. *Scholarpedia*, 9(1): 1463, 2014.

F. Ducatelle, A. Förster, G.A. Di Caro, , and L.-M. Gambardella. New task allocation methods for robotic swarms. In *Proceedings of the* 9^{th} *IEEE/RAS Conference on Autonomous Robot Systems and Competitions*. IPCB-Instituto Politacnico de Castelo Branco, 2009.

G. Dudek and M. Jenkin. *Computational Principles of Mobile Robotics*. Cambridge University Press, April 2000. ISBN 0521568765.

E. Ferrante, A.E. Turgut, E. Dué nez Guzmán, M. Dorigo, and T. Wenseleers. Evolution of self-organized task specialization in robot swarms. *PLoS Computational Biology*, 11(8):1–2–1, 2015.

O. Gigliotta, M. Mirolli, and S. Nolfi. Communication based dynamic role allocation in a group of homogeneous robots. *Neural Computing*, 13(3):391–402, 2014.

D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, MA, 1989.

D. Gordon. Dynamics of task-switching in harvester ants. *Animal Behaviour*, 38: 194–204, 1989.

D. Gordon. The organisation of work in social insects. *Nature*, 380:121–124, 1996.

A.J. Ijspeert, A. Martinoli, A. Billard, and L.M. Gambardella. Collaboration through the exploitation of local interactions in autonomous collective robotics: The stick pulling experiment. *Autonomous Robots*, 11(2):149–171, 2001.

T.H Labella, M. Dorigo, and J.-L. Deneubourg. Division of labour in a group of robots inspired by ants' foraging behaviour. *ACM Transactions on Autonomous and Adaptive System*, 1(1):4–25, 2006.

J. Lehman and K.O. Stanley. Exploiting open-endedness to solve problems through the search for novelty. In S. Bullock, J. Noble, R. Watson, and M.A. Bedau, editors, *Proceedings of the* 11^{th} *International Conference on the Simulation and Synthesis of Living Systems*, pages 329–336. MIT Press, Cambridge, MA, 2008.

J-B. Mouret. Novelty-based multiobjectivization. In *New horizons in evolutionary robotics*, pages 139–154. Springer, 2011.

G. Nitschke, M. Schut, and A. Eiben. Emergent specialization in biologically inspired collective behavior systems. In *Intelligent complex adaptive systems*, pages 100–140. IGI, New York, 2007.

S. Nolfi and D. Floreano. *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-organising Machine*. MIT Press, Cambridge, MA, 2001.

S. Nolfi and O. Gigliotta. Evorobot*: a tool for running experiments on the evolution of communication. In S. Nolfi and M. Mirolli, editors, *Evolution of Communication and Language in Embodied Agents*, pages 297–301. Berlin-Heidelberg: Springer-Verlag, 2010.

R.E. Page. The evolution of insects societies. *Endeavour*, 21(7):114–120, 1997.

R.E. Page and S.D. Mitchell. Self-organisation and the evolution of division of labour. *Apidologie*, 29:171–190, 1998.

G. Pini, A. Brutschy, M. Frison, A. Roli, M. Dorigo, and M. Birattari. Task partitioning in swarms of robots: An adaptive method for strategy selection. *Swarm Intelligence*, 5(3–4):283–304, 2011.

G. Pini, A. Brutschy, C. Pinciroli, M. Dorigo, and M. Birattari. Autonomous task partitioning in robot foraging: An approach based on cost estimation. *Adaptive Behavior*, 21(2):118–136, 2013a.

G. Pini, M. Gagliolo, A. Brutschy, M. Dorigo, and M. Birattari. Task partitioning in a robot swarm: A study on the effect of communication. *Swarm Intelligence*, 7(2–3):173–199, 2013b.

M. Quinn. A comparison of approaches to the evolution of homogeneous multi/robot teams. In J.H. Kim, B.T. Zhang, G. Fogel, and I. Kusku, editors, *Proceedings of them International Conference on Evolutionary Computation*, volume 1, pages 128–135. IEEE Press, 2001.

T.D. Seeley. Social foraging in honey bees: how nectar foragers assess their colony's nutritional status. *Behavioral Ecology and Sociobiology*, 24(3):181–199, 1989.

K.O. Stanley and R. Miikkulainen. Competitive coevolution through evolutionary complexification. *Journal Artificial Intelligence Research*, 21:63–100, 2004.

V. Trianni. Evolutionary robotics: model or design. *Frontiers in Robotics and AI*, 1:1–6, 2014.

V. Trianni and S. Nolfi. Engineering the evolution of self-organizing behaviors in swarm robotics: A case study. *Artificial Life*, 17(3):183–202, 2011.

E. Tuci. Evolutionary swarm robotics: genetic diversity, task-allocation and task-switching behavior. In M. Dorigo, M. Birattari, S. Garnier, H. Hamann, M. Montes de Oca, C. Solnon, and T. Stützle, editors, *Proceedings of the* 9^{th} *International Conference on Swarm Intelligence*, volume 8667 of *LNCS*, pages 98–109. Springer, 2014.

E. Tuci and V. Trianni. On the evolution of homogeneous multi-robot teams: Clonal versus aclonal approach. In T. Ziemke, C. Balkenius, and J. Hallam, editors, *Proceedings of the* 12^{th} *International Conference on Simulation of Adaptive Behavior*, volume 7426 of *LNCS*, pages 391–400. Springer, 2012.

E. Tuci and V. Trianni. On the evolution of homogeneous multi-robot teams: Clonal versus aclonal approaches. *Neural Computing and Applications*, 25(5): 1063–1076, 2014.

M. Waibel, L. Keller, and D. Floreano. Genetic team composition and level of selection in the evolution of cooperation. *IEEE Transaction of Evolutionary Computation*, 13(3):648–660, 2009.